The University of Chicago

AN EXPERIMENT WITH MEASUREMENT SCALES IN A CURRICULUM UNIT OF WORLD-CITIZENSHIP

A PART OF A DISSERTATION SUBMITTED TO THE FACULTY OF THE DIVINITY SCHOOL IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PRACTICAL THEOLOGY

1932

By

HERMAN STUART KEITER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1934

The University of Chicago

AN EXPERIMENT WITH MEASUREMENT SCALES IN A CURRICULUM UNIT OF WORLD-CITIZENSHIP

A PART OF A DISSERTATION SUBMITTED TO THE FACULTY OF THE DIVINITY SCHOOL IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PRACTICAL THEOLOGY
1932

By
HERMAN STUART KEITER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS
1934

R. UNIV.
BIBLIOTHEEK
LEIDEN

I PROBLEM AND PROCEDURE

Attempts at testing or measuring the results of education are as old as the first oral quiz, and probably as old as education itself. But attempts at objective testing, measurement in which the influence of the teacher's prejudices and fallibility of judgment are lessened, have come late in educational history. Objective testing appeared first on matters of fact. Bible knowledge tests still maintain their popularity. Gradually matters of opinion came to be included, and with the Character Education Inquiry of Hartshorne and May the testing movement boldly occupied the kingdom of moral education.[1]

Parallel in time to this development of objective testing has appeared a changing attitude toward things taught--the curriculum. Half a century ago curricular material seemed fairly well settled. Teachers and scholars preeminent in their fields produced combinations of fact and method known as textbooks, and around these textbooks centered the procedure of the class. But as learning came to be conceived more and more as centering around experience, the single textbook, hitherto autocrat of the classroom, found his throne become more and more unstable. A variety of pretenders clamored for it. Two texts, one for the teacher and another for the pupil, solve the problem, said some authorities. A source book for teacher and pupil, a pupil's notebook, a reference library--all found supporters.

---

[1] "In the judgment of an eminent psychologist, these two investigators, Hartshorne and May, made more progress in the measurement of human conduct in three years than was made in ten years by the earlier students of the measurement of intelligence." Nevin C. Harner, "Is Religious Education to Become a Science?" Religious Education, March, 1932, pp. 203-204.

A few years ago these two movements found a common cause, small, unobtrusive enough, but a cause for which they could join forces. The measuring of character, the latest refinement of testing, was accomplished by analytical instruments which sampled various forms of conduct, which took specimens of various parts of experience in order to judge the whole. And this very nature of measuring instruments, the "sampling of various forms of conduct," led to their use as teaching instruments apart from their regular measuring use. That is, measurement tests have been used in such ways as stimulating interest, starting discussion, opening avenues to facts, and keeping groups from wandering far afield--all quondam duties of the textbook.[1]

Here the reason for this experiment came into being. Is such teaching use of a measuring instrument practicable? Does it result in increased learning? Curricular materials should begin, continue, and end in experience, authorities agree. What more appropriate, then, for curricular use than measuring instruments, samples of experience objectively selected, arranged, interpreted? Is it possible that a time will come when a measuring instrument prepared for teaching will be a prerequisite to any up-to-the-minute curriculum plan or unit of teaching? May the measuring instrument aspire to the throne of the textbook?

---

[1]Such use has not been extensive. Full description of most past teaching uses of measurement tests are included in the following: G.B. Watson, The Measurement of Fairmindedness (New York: Teachers College, Columbia University, 1925); H.C. Morrison, The Practice of Teaching in the Secondary Schools (Chicago: University of Chicago Press, 1926), Chap. IX; B.S. Winchester, Young Peoples Studies (Boston: Pilgrim Press, 1931). An instrument planned for such use is the "Attitude Test of World Relations" constructed by E.J. Chave of the University Of Chicago in cooperation with the Bureau of Research of the International Council of Religious Education for use with Lobingier's experimental unit, How Big Is Your World?

The Problem Limited and Defined.--It was obviously not within the province of this experiment to prophecy concerning the future of teaching with measurement scales. But it could attack the problem: Do measurement scales possess definite teaching value? This experiment attempted to locate and evaluate any possible teaching use of such scales. The teaching was of a curriculum unit of world citizenship to a group of Junior High School church school learners. Measuring scales are a refinement of the more common types of measuring instruments; they provide objectivity and make possible an exactness of marking by allowing the subject to mark anywhere along a line.[1] They can be said to possess teaching value if their use as teaching instruments results in a noticeable improvement in group learning over similar teaching without the instruments.

The Set-Up.--Possible teaching value in measurement scales was investigated in this way: Two groups, so selected as to be equivalent in important characteristics, were taught World Citizenship in the same way except that with one the scale was used as a teaching instrument. Comparison of the resultant learning in the two groups offered an index to the teaching value of the scales.

Checks on This Method.--Two pairs of groups were used instead of one, and the Rotation technique was added to the Equated-Group technique, in this fashion:

1. Five Junior High School groups were selected on a basis of equation determined by these factors--
   World Citizenship opinion--according to a composite score of the Scale Used in Teaching and the Chave-Lobingier Test.

---

[1]Thurstone and Chave, The Measurement of Attitude (Chicago: University of Chicago Press, 1929).

intelligence--by the Terman Group test.
socio-economic status--by the Sims test.
age, attendance, and school grade.

2. One group, C, the control group, had regular church school work. It was measured before and after the experiment to check on any possible shift of community opinion.

3. The other two pairs, X and Y, and R and S, were taught as described above. The teaching of R and S, which acted as a check upon the teaching of X and Y, was done under less favorable conditions, with restricted time, a minimum of activity and appreciational phases, and with a teacher less acquainted with the material.

4. The Rotation Technique. Group X was taught without the aid of the scales and Group Y with the aid of the scales in the race phase of world citizenship. Then in the teaching of the peace phase the arrangement was reversed: Group Y was taught without the aid of the scales, and Group X with the aid of the scales. The teaching of Groups R and S was similarly arranged.

5. Thus were four situations set up (nomenclature thus: first letter, name of group; next two letters, st, scale-taught, and ns, non-scale-taught; final letter, type of experience, either race, r, or peace, p):

Situation I: Ystr, Xnsr
Situation II: Xstp, Ynsp
Situation III: Sstr, Rnsr
Situation IV: Rstp, Snsp

Controlled Conditions.--The following conditions were kept the same in the groups to be compared:

1. Teaching time--same in groups compared.
2. Teachers--same in groups compared.
3. Intelligence, social background, world citizenship opinion, age, attendance, and school grade were made approximately equivalent in the groups compared by means of the equating process.
4. Possible difference in groups were eliminated by rotation.[1]
5. Check on possible opinion change in community given by Group C, the control group.

Certain conditions could not be kept entirely the same in the groups to be compared: attendance, home opinion, special point of view such as of R.O.T.C. members, and adolescent love affairs.

---

[1]See item 4 above.

Measuring Instruments: 1. Chave-Lobingier Test.--This consists of fifty statements on race, peace, and missions to be marked true or false. It was used in two ways: First, to help in equating the groups to be compared; and, second, to help check on group learning by measuring opinion before and after the experiment.[1]

2. Scales Used in Teaching: Triple-Section Instrument Developed by Investigator.--This instrument was used in the same way as the Chave-Lobingier Test, but with this additional all-important use: It was employed in the actual "teaching with scales," the process whose results were measured in this experiment. There was constructed a complete set of three sections on both race and peace phases of World Citizenship, or Interpenetration of Culture. Samples follow:

---

[1]For this second use, fifteen race and fifteen peace statements were selected from the fifty. Samples: 5. The average negro is mentally as capable of education as the average white.
13. National honor should be defended at all costs.

I Opinion Section, "What Do You Think?" Opportunity is given the subject to register opinion of any degree (even prejudice) along certain lines. An example:

VIII How do you feel about pacifists, persons who declare they are patriots but say they serve their countries best by refusing to be soldiers or even to help in a war? Vote by placing an "X" anywhere along the line.

Put an X
on this line
-- Pacifists are all either cowards or enemy spies.

-- I think pacifists cannot be patriots, though they may be sincere.

-- I am undecided about pacifists.

-- I think pacifists are just as good patriots as militarists, especially since the Kellogg treaty.

-- Pacifism is the only sensible way of fighting war.

II Action Section, "What Would You Do?" consisting of multiple-choice situations. A sample:

3. Suppose the school authorities allowed Negro pupils to swim in the same swimming pool as the white children, and some of your friends said they would not swim there if the Negroes were allowed to. What would you do?

( ) I would tell my friends they were foolish to object, that Negroes were just the same as whites except for the color of their skin--and that didn't matter. And I would be quite willing to swim there.
( ) I would not swim there, because my parents would not let me.
( ) I would not swim in the same water Negroes swim in; I could not enjoy it at all.
( ) I would certainly protest and get my parents to do the same.
( ) _____

III Stimulation Section, "How Well Can You Guess?" consisting of Yes-No and Completion statements of fact; not used for measuring, though appearing as if for such use. Samples:

In the blank spaces write the number you think is correct:

4. During its history, the United States government has submitted _____ disputes to arbitration.

5. _____ arbitration decisions (on U.S. cases) were not accepted by the disputing nations.

Draw a circle around "yes" if you think the statement true; "no" if untrue.

Yes No 1. It is better to eat a man without cooking him than to cook a man without eating him.

Yes No 4. White men are in certain ways, such as hair structure, more like apes than are Negroes.[1]

Developing the Scale Used in Teaching.--This instrument was perfected by repeated experimentation and revision. It was tried out on group after group in its experimental stages. From data furnished by these experimental markings, each scale or problem was evaluated on the basis of these three criteria, and was eliminated or changed to comply if it was deficient in any respect:

1. Do the markings taken together display normal distribution? (Not all marking the same item!)
2. Does the same person mark the same way at different times?
3. What evidence is there of misunderstanding or ambiguity?

Scoring of the Scale.--After the race instrument was thus perfected, a numerical value or score for each item of each scale was thus assigned: five graduate students rated each item on a sliding scale of one to eleven in race friendship, in the race scales.[2] These ratings were checked by the distribution of the experimental markings to arrive at a final scoring value.

---

[1]Correct answers: 243 disputes (the usual guess is 3 or 4), 1 unacceptable decision; Yes on both true-false statements (the former a contrast between cannibalism and lynching).

[2]The peace Scale Used in Teaching was used only for teaching and was not thus rated.

Teaching Units.--Since the groups were to be similarly taught except for the scales, teaching plans had to be prepared. A unit of learning (or teaching) is always a unit of experience. Sudden withdrawing of the hand from a hot stove is a simple unit of experience (one response to one stimulus), is usually a unit of learning, and may be a unit of teaching. Most units include response to combinations of stimuli, or "situations." And these situations in adult life often include vast areas of human experience (numberless stimuli and responses), as in a world friendship situation or unit.

Units developed.--The teaching units prepared for this experiment attempted to preserve as clearly as possible the one-stimulus-one-response typical element and never to vary from the situation-response type, while including a conveniently large area of experience. Each of the five outlined was intended to develop an integral type of experience in the large field called "World Citizenship" or "Interpenetration of Culture." The five planned were:

1. Is the race problem important enough to study?
2. How should we treat persons of different races and nationalities?
    a. Are they inferior?
    b. Where does race feeling come from?
    c. Can Christianity throw any light on the problem?
3. How do we want our nation to treat nations of other races?
4. Why do many people consider the problem of war of great importance? Shall we study it?
5. What do we want to do about war?

Process of Development.--A search was made for suggestions from denominational and non-denominational curricula. The general outline of the units followed the thirteen steps of learning as developed by Bower.[1] Appropriate source material to be made

---

[1] Character Through Creative Experience (Chicago: University of Chicago Press, 1930), Chap. VII.

available to the learners was gathered in the form of books discussing the problem, and in additional primary sources such as Negro newspapers and magazines for the units on the race problem. Appreciational and action phases were worked out with the groups, as, for example, selecting bits of Negro poetry for use in the worship service, forming a committee and through it interviewing the pastor on the possibility of a Negro group visiting the church, investigating the possibilities of an international language and of international correspondence, or of talking with a Negro acquaintance and bringing to the group a report of the Negro's opinion on the importance of the race problem. Notebooks to be used at the discretion of the learner were provided for each member of all groups except the control group. A copy of the scales was included in the notebooks of the learners when they were taught with the aid of the scales, but were not included when the learners were taught without the aid of the scales.

II THE TEACHING VALUE OF MEASUREMENT SCALES

This experiment exhibits one major finding: Measurement scales do possess definite value as teaching instruments--at least scales of this type, in a type of experience similar to World Citizenship, in a teaching situation with similar teachers and learners of approximately the same age and background. That is, no claim can be made that all measurement scales have such teaching value in all situations. But the conclusion can reasonably be extended to include most cases of similar scales in similar teaching with college-trained teachers and learners no younger than Junior High age.

Data from tests taken before and after the experiment and from personal interviews with the subjects support this conclusion, that in a situation similar to the above, teaching, even of the "experience" type, can be improved by the teaching use of measurement scales. And descriptive data attempt to locate the specific teaching values, to point out at what steps or phases in the learning process these values arise.

Of course, the results of this experiment do not begin to answer all the problems in the teaching use of measurement scales. The report must be "not proven" even upon several questions specifically undertaken by this project.[1] But several results can be and are recorded with supporting data.

---

[1]Results were negative in these sub-investigations: How measure prejudice so it can be avoided? Does use of a measurement instrument for teaching lessen measuring value? Does use for measuring lessen teaching value? How does teaching value vary with differing teaching conditions?

A. Scales Do Have Teaching Value

The learning of scale-taught groups was found to be greater than that of non-scale-taught groups in all of the four situations:

Situation I:	Ystr,	Xnsr[1]
Situation II:	Ynsp,	Xstp
Situation III:	Sstr,	Rnsr
Situation IV:	Snsp,	Rstp

Learning was measured statistically by tests taken before and after the teaching. They showed the scale-taught group superior in every situation except Situation III, where the superiority of the scale-taught group was so slight as to be negligible.

Learning was also measured by the response of the pupils to selected questions in private interviews. The answers of the pupils showed the scale-taught group superior in learning in all four situations.

The interviews also included questions on the pupil's attitude toward the scales. A great majority of the pupils commended them highly and recommended their future use.

1. The Testimony of the Interviews

Each pupil was interviewed privately, and asked each question in a non-committal way. The answer was recorded immediately by a stenographer.

The Opinion of the Learners on the Scales.--Out of the 32 learners, only 2 disliked the scales, 6 made no decisive response, 8 liked them, and 15 were enthusiastic about them. A large majority favored having similar scales in a future study, and offered to help construct them. More than half declared specifically that the scales helped them to "see both sides of the question." The value of this response was heightened by the

---

[1]For explanation of nomenclature, see p. 4.

fact that for the purposes of the experiment it was necessary that the learners mark five tests: Once for the Terman Group Test, and twice each for the Chave-Lobingier Test and the Scale Used in Teaching. And naturally the learners were a bit weary with what seemed to them too much marking of tests.

Complete responses follow:

THE OPINIONS OF THE LEARNERS ON THE TEACHING VALUES OF THE SCALES

First Question.--What do you feel about using these question blanks?--the mimeographed ones in your notebooks particularly, not the printed ones. Were they most interesting, fairly good, easy to get tired of, or unpleasant? Were they too much like school? Do you feel they helped you to think about different sides of the problem?

Group X

1*-Very interesting, I liked them a lot. They help to give sides to problems because they have various degrees of questions at the top and bottom, that is, the top is one extreme and the bottom another.

2-All right.

3-I think it helps improve your thoughts on the subject. Part of it is on one side, part on the other.

4-I think they are interesting. Give a person a different feeling toward races.

5-Well, gave you ideas of possibilities; not too much like school.

6-A good way of finding how people think about a problem; it may change an opinion because it mentions some conditions people may not have thought of before.

7-I like them; I like to work them.

8-

Group Y

11-A lot of fun, just to see what answers would come out.

12-Worthwhile, not too much like school; shows both sides of question, shows how to act.

13-Fairly interesting, made me think. Those blanks knew what I'd do; made me stop and think as I looked over the situation.

14-Not hard, just took a little thinking.

15-They were swell; I would like to have them over again.

16-Interesting, made me think about it; they aren't too much like school.

---

*Numbers refer to various learners. Group X consisted of learners numbered from 1 to 8; Group Y from 11 to 18; Group R from 21 to 28; Group S from 121 to 128; Control Group the "tens" from 10 to 80.

Group Y (cont.)

17-I think they are perfectly all right. They get monotonous but I think they are a good thing.

18-Have very good examples. They aren't too much like school. Give both sides of the problem.

Group R

21-

22-Interesting. Like school in a way but not much. Give very different views.

23-

24-Fine. Not like school, just O.K. Present different views of the problem.

25-They are all right. I don't think they are much like school. Some give different ideas--I didn't get a chance to read all of them.

26-Rather interesting--not much like school. Give different views of question.

27-Not too much like school.

28-Interesting; tire easy. Too much like school? Yes, ma'am.

Group S

121-Find them all right; if you have different ideas it helps you form more of an opinion. I like them.

122-I really think they are very interesting, they aren't much like school. Help one to get different views.

123-Yes, quite interesting. Don't have anything like it in school.

124-

125-Not tiresome, not like school; gave some different points of view.

126-I like them; they aren't too much like school.

127-I don't know; rather like school.

128-I like them, they aren't tiresome or too much like school. Give different views.

SUMMARY.--Feeling of the Group on the "Question Blank."

Enthusiastic about their use	15
In favor	8
Doubtful	3
Against	2
Total responding	28

Second Question.--In another study do you think it would be interesting to have these for those who wanted them?

Group X

1-Carrying out the idea farther? I think it would.

2-I think so.

3-It would be interesting for anybody.

4-Yes.

5-Aviation.

6-Yes.

7-I do think they would.

8-

Group Y

11-Yes, if you leave out the tests.
12-I don't know.
13-Might be very interesting.
14-Yes.
15-Sure.
16-I think it would because this one was really interesting.
17-Yes.
18-Yes.

Group R

21-
22-In a way.
23-
24-Sure.
25-Yes.
26-I do not know.
27-No.
28-Sports and other activities, band.

Group S

121-Yes.
122-I think they would be interesting, what other courses would there be?
123-No.
124-
125-Don't know.
126-Yes.
127-I think it would.
128-I think it would.

SUMMARY.--Feeling of the group on using similar scales in another study.

Enthusiastically in favor	7
In favor	15
Doubtful	4
Against	2
Total responding	28

Third Question.--Do you think you would like to help make such question blanks?

Group X

1-Well, I don't know if I'd be of much help. If I gave it some thought perhaps I could.
2-I might.
3-I would like to help.
4-Yes, if no writing. I don't like writing.
5-I wouldn't know what to put in.
6-If I could be of any.
7-Uhhuh.
8-

Group Y

11-I don't know.
12-I don't know, not sure.
13-Yes.
14-(hesitantly) Yes.
15-I would say "yes," that would suit me to the dot.
16-Yeh.
17-Oh, wait a minute, hard to tell, I would not mind helping if I could find time.
18-I don't think I could.

Group R

21-
22-Don't think I would do that.
23-
24-Yes.
25-Yes.
26-I don't know.
27-No.
28-No.

Group S

121-Uhhuh.
122-I'd like to.
123-I don't know.
124-
125-No, I don't think so.
126-Maybe.
127-Uhhuh.
128-Yes, I would.

SUMMARY.--Feeling of the group on helping to make such scales.

Enthusiastically willing	4
Willing	15
Doubtful	8
Unwilling	3
Total Responding	28

The Interviews Show Scale-Taught Groups Superior in Learning.--The answers to the following seven interview questions on learning, particularly to the last four, furnish probably the most important data of the experiment. These responses show a marked superiority in learning of the scale-taught groups over the non-scale-taught groups.

The first three questions asked the learners their opinion of the study, its importance, its value as compared to other studies, and any new ideas they received from it. Direct answers to the questions, it was realized, would be worth little or nothing,

but it was hoped that the pupil would add some comment of his own. The effect hoped for was the same as if the question had been asked, "What comes into your mind first of all when these studies are mentioned?"

The answers, while interesting, did not show a great deal of difference between the scale-taught and non-scale-taught groups. Out of twelve indices (the answers to each of the three questions applied to each of the four situations), one showed a slight inferiority in learning of a scale-taught group, three showed slight superiority,[1] and the other eight showed the groups practically the same so far as learning was concerned. Of course, there was little or no possibility of manipulation by the questioner in these free expression questions, so that the results must be given more weight than might first appear.

Detailed answers follow:

INTERVIEW, SECTION I. GENERAL IMPRESSION (AN EFFORT TO GET FREE EXPRESSION ON THE SCALES AND THEIR EFFECTS).

First question. Have the things we have been talking about seemed important?

Second question. Do you think the programs were worthwhile?--in comparison with anything you have had before.

Reactions to these first two questions seem not significant enough to be worth quoting. The only possible conclusions which may be drawn: first, the learners do not seem particularly enthusiastic; second, in Situations III and IV, the answers to the first question seem to show the scale-taught group superior in race learning, but inferior in peace learning.

Third question. Did you get any new ideas that you can recall?

---

[1] An attempt to objectify the terms "slight," "moderate," and "strong," as applied to superiority is described below, p. 19.

Answers by Groups X and Y, learners in
Situation I, Ystr, Xnsr, and in
Situation II, Ynsp, Xstp

Group X (nsr, stp)

1*-Well, those questions last Sunday, about the number of people killed because of the war. How New York might be blotted out in a few hours, in the number of disputes settled by arbitration, in the speed of tanks. I enjoyed looking for articles in newspapers, magazines, etc.

2-About the association of blacks and whites, to treat them as human beings.

3-Sure, well, that almost all other races are treated almost the same as colored.

4-Well, let's see, I got an idea on how the colored and white races differ, that the colored have brains as much as we do. Ideas I never stopped to think about, I never thought they would have newspapers and magazines, that they might be our equals, that they might have professions.

5-Suppose so, but my opinions have not changed.

6-That we should not feel superior to other nations.

7-I don't know.

8-I don't know, about other races, treating them so low.

Group Y (str, nsp)

11-Not any I can think of now.

12-I think so but I can't recall them now.

13-No, it seemed to make things a bit clearer and bring them to mind. I always thought of it but paid no attention, this brought it up and made me think about it.

14-Yes, that the race problem is bigger than I thought it was before, how other people of other nations look on us and how we look on them, what it means to us and to them.

15-Yes, not to begrudge, not to call names, not to be afraid, to treat others as brothers as they are, to treat them kindly.

16-Never had any ideas about it before, I never thought of it before.

17-I can't think of them now, I suppose I have.

18-That you shouldn't think of other people as lower than yourself, but think of them as just as good if not better in some ways.

CONCLUSION SITUATION I--Learning in the scale-taught group, Y (str) seems about equal to that in Group X (nsr).

"Race" Answers (Situation I) to
Interview Section I Third Question

	Group Y (str)	Group X (nsr)
No	3	1
?	0	3
Yes	2	1
Definite Instances	3	3

See p. 12, N..

CONCLUSION SITUATION II--The scale-taught group (Xstp) seems to show a slight superiority in peace learning, since the one answer from the peace angle (and a very complete answer) was from the peace-scale-taught group.

Answers by Groups R and S, learners in Situation III, Sstr, Rnsr, and in Situation IV, Snsp, Rstp

Group S (str, nsp)

121-Yes, about the attitude we should take toward the colored, and stuff like that.

122-Different opinion of negro. Before I thought that they were lower than we were, but now I think there is not much difference.

123-Yes, I learned some things I never knew before.

124-

125-I don't know.

126-Yes, about the negro--I feel a real difference toward them. I think the negro should be treated as well as we are.

127-Things about the next war, about the poison and stuff; if all stick together there will be no more war. I don't know, just a lot of stuff.

128-Yes, in a way. I don't know.

Group R (stp, nsr)

21-

22-Well, I know something about what the next war will be like if we have one.

23-

24-Yes, how to treat a negro.

25-

26-I think I did, yes, very much. There is not much difference in mental ability between races; about the same if under the same natural conditons.

27-Can't remember any.

28-Yes, I learned more about colored people and we see how they live.

CONCLUSION SITUATION III--The scale-taught group in the race units seems to show slightly superior learning according to the tabulation:

"Race" Answers (Situation III) to Interview Section I Third Question

	Group S (str)	Group R (nsr)
No	0	0
?	3	5
Yes	2	0
Definite Instances	3	3

CONCLUSION SITUATION IV--The learning of the groups on the peace units seems about the same since one in each group mentions it.

This last question of those calculated to stimulate free expression seems the only one to have fulfilled its purpose. The results of this first type of question, inasmuch as the expression is free, are consequently of proportionally more weight than the others.

An Attempt to Define Various Differences in Learning.--The tabulation immediately above suggests a method of obtaining an objective check upon terms applied to the amount of superiority in learning of one group over another. In this case (Interview Section I Question Three as applied to Situation III, Sstr and Rnsr) the term "slight" superiority, as may be noticed, is applied to Group X because two persons out of the eight answered "Yes," that they did get new ideas (though without specifying) rather than being undecided--the other six have the same types of responses, group compared to group.

If we consider "No," "Yes," and "Definite Instances" as class intervals dividing the persons according to responses, then we can judge the superiority of a group by the number of class interval shifts necessary in the inferior group to bring it to equality with the superior group.[1]

Types of Superiority.--On this basis, where the groups are so nearly the same in response that a shift of one person one class interval will make them the same, the difference between the groups is considered negligible.

"Slight" superiority could then be defined as a situation in which the inferior group must be lifted from two to four class interval shifts, as in Situation III above, and "Moderate" superiority as a situation in which the inferior group must be lifted

---

[1] A "Class interval shift" would be defined as the moving of one person one class interval in the tabulation.

from five to seven class interval shifts. "Strong" superiority would thus be eight class interval shifts or more--an average of at least one class interval shift per person. Eleven class interval shifts or more would indicate "very strong superiority."

The use of these arithmetical class interval shifts must not be understood as a claim to mathematical accuracy. It is merely an attempt to get an objective check upon what must essentially be a subjective judgment. And the quality of the incidents cited may often change considerably that mathematical index. For example, in the response of Groups X and Y to Interview Section II Question Three, the scale-taught group appears only negligibly superior on the basis of class interval shifts.[1] But the quality of the definite instances cited are so much higher in the scale-taught group that it can be fairly considered moderately superior in response to this question.

Weightiest Data of the Experiment.--The second section of the interview attempted to probe the learning of the subjects by direct questions concerning changed action. Most convincing returns came from these four questions:

> 1. Have you run into any problems in which this study has led you to act differently?
>
> 2. Has this course in any way made you feel or act differently, or do you think you have behaved just about the same as you did before?

---

[1]This "class interval shift" yardstick obviously falls far short of complete accuracy, since it assumes that the difference between a "No" answer and a "?" answer is the same as the distance between a "Yes" answer and a "Definite Instance" answer. And that assumption is often inaccurate--so the yardstick's inches are not always the same. Its use seems justified for this reason: Two persons comparing the answers of two groups might not agree as to which indicated superior learning. But they would agree upon the tabulation as given by this yardstick.

3. Have you talked over any of these things with the folks at home or with your chums?--what happened?

4. Have you read anything on the subject besides the books, magazines, and newspapers provided? What of your reading did you like best?

Each answer to these questions applied to two situations for each pair of groups. For example, the answers of Group X and Y were compared with each question. The pupil may have answered either from the peace angle or from the race angle, or possibly both. So the race answers of the two groups were compared, and the peace answers likewise.[1] Group X was taught with the scales in the Peace Units (Situation II, Ynsp, Sstp), so if the Peace Answers of Group X were superior to those of Group Y, this one index would show teaching value in the measurement scale. A similar index could be determined for the Race Teaching where Group Y was scale-taught (Situation I, Ystr, Xnsr). And similarly for Groups R and S (Situation III and IV). Thus each question would furnish one index for each situation or a total of sixteen indices for these four questions.

Out of the sixteen indices not one showed a non-scale-taught group superior to its scale-taught mate. Six indices showed no difference or only a negligible one between the two types of teaching, and the other ten showed the scale-taught group superior to the non-scale-taught, three indices showing them slightly superior, five moderately, and two strongly.[2]

Detailed results follow:

---

[1]The Peace Unit was taught after the race unit but so much less time could be given to it that most of the answers were from the race angle.

[2]TABLE I, p. 23, summarizes these differences for both Sections I and II of the interviews.

INTERVIEW, SECTION II. DIRECT QUESTIONS ON POSSIBLE CHANGED ACTION OF THE SUBJECTS AS A RESULT OF THE TEACHING.

First Question. Have you run into any problems in which this study has led you to act differently?

Answers by Groups X and Y, learners in
Situation I: Ystr, Xnsr, and in
Situation II: Ynsp, Xstp

Group X (nsr, stp)

1*-I think more carefully about swimming in the same pool with negroes but I haven't yet reached a conclusion. I think even more about questions of war and peace, and about settling disputes without war.

2-No.

3-I do.

4-I don't think so.

5-I have not taken swimming because of the negroes who are allowed to use the pool. I haven't started to since the program, I have thought about it, but I don't. My idea is probably silly for the negroes are probably as clean as those whites, but I don't.

6-No.

7-I think it is no harm to go swimming with a nigger.

8-I don't think so. Oh, our white maid wouldn't eat with our colored maid, so I spoke to my mother about it and now they eat together.

Group Y (str, nsp)

11-Well, I haven't yet but if I did I know I'd act differently.

12-No.

13-I don't know. There is a negro in the study hour that I like and he has many white friends.

14-Yes, the way you treat a foreign boy in school and outside. We were playing golf and there were only three of us so we had to fill in with a colored guy. We played just as if he was white.

15-A colored person came up to me asking about a certain street and I told him, I didn't before.

16-All are new.

17-Well, we have colored girls in our class and I don't think I have been quite so high hat; I am not certain it was the result of this, but it may have been.

18-There are a lot of colored kids around. Some of the fellows won't have anything to do with them. I never minded but I feel better in being friendly to them now.

See p. 12, N.

TABLE I

THE TEACHING VALUE OF MEASUREMENT SCALES IN RACE AND PEACE TEACHING AS SHOWN BY PERSONAL INTERVIEWS WITH THE SUBJECTS--SUPERIORITY OR INFERIORITY IN LEARNING OF SCALE-TAUGHT GROUPS AS COMPARED TO EQUIVALENT NON-SCALE-TAUGHT GROUPS AS SHOWN BY RESPONSES IN THESE INTERVIEWS.

Interview	Superiority or Inferiority in Learning of Scale-Taught Groups as Compared to equivalent Non-Scale-Taught Groups.			
	Groups X and Y		Groups R and S	
Sec. I	Situation I Ystr, Xnsr	Sit. II Ynsp, Xstp	Sit. III Sstr, Rnsr	Sit. IV Snsp, Rstp
Question One*	No Difference	No Difference	Slight Superiority	Slight Inferiority
Two	No Difference	No Difference	No Difference	No Difference
Three	No Difference	Slight Superiority	Slight Superiority	No Difference
Sec. II				
Question One	Moderate Superiority	Slight Superiority	Moderate Superiority	No Difference
Two	No Difference	No Difference	Moderate Superiority	No Difference
Three	Moderate Superiority	No Difference	Strong Superiority	Slight Superiority
Four	Moderate Superiority	No Difference	Moderate Superiority	Slight Superiority

*Questions given above, pp. 16, 17, 20, 21.

SUMMARY.--SUPERIORITY OR INFERIORITY IN LEARNING OF SCALE-TAUGHT GROUPS AS COMPARED TO EQUIVALENT NON-SCALE-TAUGHT GROUPS: SLIGHT INFERIORITY-1; NEGLIGIBLE OR NO DIFFERENCE-14; SLIGHT SUPERIORITY-6; MODERATE SUPERIORITY-6; STRONG SUPERIORITY-1.

CONCLUSION SITUATION I--IN THE "RACE" UNITS THE SCALE-TAUGHT GROUP SHOWS A DEFINITE SUPERIORITY OF MODERATE AMOUNT, ACCORDING TO THE TABULATION (SEVEN CLASS INTERVAL SHIFTS):

"Race" Answers (Situation I) to Interview Section II Question One

	Group Y (str)	Group X (nsr)
No	1	3
?	1	2
Yes	2	1
Definite Instances	4	2

CONCLUSION SITUATION II--THE RESPONSES GIVING AN INDEX TO THE PEACE LEARNING ALSO SHOW A SLIGHT SUPERIORITY FOR THE SCALE-TAUGHT GROUP, SINCE ONE IN THAT GROUP EMPHASIZES THE SUPERIOR IMPORTANCE OF THE WAR-PEACE QUESTION, WHILE IT IS NOT MENTIONED AT ALL IN GROUP Y (nsp).

Answers by Groups R and S, learners in Situation III, Rnsr, Sstr, and in Situation IV, Rstp, Snsp

Group S (str, nsp)

121-No.
122-There is a negro girl in my dramatic club at school. I sit next to her and speak to her; another girl asked me how I could do it. I told her that the only difference was one of color. Still she said she couldn't see how I could.
123-
124-
125-Don't think so.
126-No.
127-We talk about it at school too.
128-No.

Group R (nsr, stp)

21-
22-How we should treat people of other races.
23-
24-I was walking down the Midway with a fellow and we saw a Negro throwing away paper. The fellow said Negroes should not be allowed on the Midway. (What did you do?) I told him they should that they were just as good as he was.
25-I don't feel quite the same. A Negro stabbed a woman I knew and she died. (How did you feel before?) I felt they were just as we are and now I feel that some aren't and some are.
26-I don't know.
27-No.
28-No.

CONCLUSION SITUATION III--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP SEEMS MODERATELY SUPERIOR TO THE OTHER IF JUDGED ON THE QUALITY OF THE INSTANCES CITED, THOUGH THE TABULATION INDICATES ONLY SLIGHT SUPERIORITY (TWO CLASS INTERVAL SHIFTS):

"Race" Answers Situation III to Interview Section II Question One

	Group S (str)	Group R (nsr)
No	3	3
?	3	4
Yes	0	0
Definite Instances	2	1

CONCLUSION SITUATION IV--PEACE PROBLEMS ARE NOT MENTIONED SO NO CONCLUSION CAN BE DRAWN.

Second Question. Has this course in any way made you feel or act differently or do you think you have behaved just about the same?

Answers by Groups X and Y, learners in
Situation I: Xnsr, Ystr, and in
Situation II: Xstp, Ynsp

Group X (nsr, stp)

1-I think I got the other fellow's view-point, negroes especially. I find that many of them think of us as we think of them.

2-No.

3-I just answered that.

4-Well, there are a lot of Jewish kids in our neighborhood and we fellows always teased them because they are not good fighters. I don't know but I believe I have a different feeling toward them.

5-No, to tell the truth.

6-Well, I don't know. I feel now we are not superior, I felt that somewhat before but not as much.

7-I don't know.

8-I act differently toward different people; I notice other races more, in our maid's case particularly.

Group Y (str, nsp)

11-Well, I would be different in my actions toward different nationalities or races. I would treat them better.

12-Not so far.

13-Always more or less knew we shouldn't call them niggers or treat them mean but I did as the crowd did. This put a new light and made me realize that I wasn't doing right.

14-I feel differently to foreigners.

15-Oh, I do. I try to act kindly, to help, sit next in school.

16-(no comparison)

17-Hard to answer, I think not.

18-No, I can't say so, I never acted cold.

CONCLUSION SITUATION I--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP SEEMS TO DIFFER SO SLIGHTLY FROM THE OTHER GROUP THAT THE RESULTS ARE NEGLIGIBLE ACCORDING TO THE FOLLOWING TABULATION (ONE CLASS INTERVAL SHIFT):

"Race" Answers Situation I to
Interview Section II Question Two

	Group Y (str)	Group X (nsr)
No	3	2
?	1	2
Yes	1	1
Definite Instances	3	3

CONCLUSION SITUATION II--PEACE PROBLEMS ARE NOT MENTIONED SO NO CONCLUSIONS CAN BE DRAWN.

Answers by Groups R and S, learners in
Situation III: Sstr, Rnsr, and in
Situation IV: Snsp, Rstp

Group S (str, nsp)

121-No.
122-I don't just know what to say about that.
123-Well, I thought I was superior to the different races but I don't feel that now.
124-
125-I play with more girls instead of reading so much.
126-I don't think I feel different. Before I had a Jewish girl friend and I still like her.
127-About the same but a little different.
128-Yes.

Group R (nsr, stp)

21-
22-I think about it more.
23-
24-On the Midway.
25-Because the woman was stabbed.
26-A little different.
27-No.
28-No.

CONCLUSION SITUATION III--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP SEEMS DEFINITELY SUPERIOR TO A MODERATE DEGREE OVER THE NON-SCALE-TAUGHT GROUP ACCORDING TO THE QUALITY OF THE INSTANCES CITED THOUGH THE FOLLOWING TABULATION SHOWS IT BUT SLIGHTLY SUPERIOR (THREE CLASS INTERVAL SHIFTS):

"Race" Answers Situation III to
Interview Section II Question Two

	Group S (str)	Group R (nsr)
No	1	3
?	3	2
Yes	2	1
Definite Instances	2	2

CONCLUSION SITUATION IV--PEACE PROBLEMS ARE NOT MENTIONED SO NO CONCLUSION CAN BE DRAWN

Third Question. Have you talked over any of these things with the folks at home or with your chums--what happened?

Answers by Groups X and Y, learners in
Situation I: Xnsr, Ystr, and in
Situation II: Xstp, Ynsp

Group X (nsr, stp)

1-I heard these things discussed at home before. We talked it over with friends and some have the same and some different points of view.

2-To my Dad; didn't say much.

3-Not at home. I told them what we were doing in League and they thought it was a good topic to work on.

4-I don't see much of my parents or brother for they work.

5-Somewhat at home; not much happened.

6-To friends about swimming but we have come to no conclusion. We still talk about it. We know we should not mind swimming in the same water with Negroes but we do.

7-No

8-To my mother; to friends to ask about the past lessons of the League.

Group Y (str, nsp)

11-At home with mother; she summed up situations.

12-No.

13-Mentioned it to mother who thought it was a good thing and that the ideas were very very helpful; to friends in class who said they were good ideas, rather interesting.

14-Home. We should treat foreign pupils as if pupils of my own country; has had a harder time than we, coming here and not acquainted. To friends, just telling each other about the statements.

15-Some, talked about it to my Dad who thinks it's important. He says he is glad I am coming. To boys who won't play with colored boys and treat them mean. They don't do it as much as they used to though they aren't perfect yet.

16-No.

17-Only told a few things we discussed here.

18-No, not to a great extent. To some friends in church; some for and some against it.

CONCLUSION SITUATION I--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP SEEMS MODERATELY SUPERIOR TO THE OTHER AS JUDGED ON THE QUALITY OF THE INCIDENTS THOUGH THE NUMERICAL TABULATION SHOWS NEGLIGIBLE DIFFERENCE:

"Race" Answers Situation I to Interview Section II Question Three

	Group Y (str)	Group X (nsr)
No	2	2
?	0	0
Yes	0	1
Definite Instances	6	5

CONCLUSION SITUATION II--PEACE DISCUSSIONS ARE NOT MENTIONED SO NO CONCLUSION CAN BE DRAWN.

Answers by Groups R and S, learners in Situation III: Rnsr, Sstr, and in Situation IV; Rstp, Snsp

Group S (str, nsp)

121-No.

122-To mother; she thinks there are no important differences. To teacher; she thought the girl was rude to say things like that (concerning negro in dramatic club).

123-To father and mother. They gave me additional views on the question.[1]

124-

125-To mother.[1]

126-To father and mother about the negro problem; some friends have the same, some different views.

127-Just talked to friends.[1]

128-Father didn't have much to say.[1]

Group R (nsr, stp)

21-

22-Discussed different views; next war, more interesting.

23-

24-Sure! Dad agreed with us.[1]

25-No.

26-Yes, thought about as I did. (What do you think?) I think the question of taking other races into our social life must be considered.

27-No.

28-No.

CONCLUSION SITUATION III--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP SEEMS STRONGLY SUPERIOR TO THE OTHER, AS JUDGED BOTH ON THE QUALITY OF THE INCIDENTS AND ACCORDING TO THE NUMERICAL TABULATION OF INDIVIDUAL ANSWERS (NINE CLASS INTERVAL SHIFTS):

"Race" Answers Situation III to Interview Section II Question Three

	Group S (str)	Group R (nsr)
No	1	3
?	1	3
Yes	1	0
Definite Instances	5	2

CONCLUSION SITUATION IV--IN "PEACE" LEARNING, THE SCALE-TAUGHT GROUP, GROUP R, SEEMED SLIGHTLY SUPERIOR TO THE OTHER, SINCE ONE LEARNER IN THAT GROUP MENTIONED A DEFINITE INSTANCE OF DISCUSSION ARISING FROM THE STUDY, WHILE IN GROUP S NO MENTION OF WAR AND PEACE WAS MADE.

---

[1]Several answers which could not be positively identified as either race or peace were classified as race, since the probabilities pointed that way. But their omission would make little change.

Fourth Question. Have you read anything on the subject besides the books, magazines, and newspapers provided? What of your reading did you like best?

Answers by Groups X and Y, learners in
Situation I: Xnsr, Ystr, and in
Situation II: Xstp, Ynsp

Group X (nsr, stp)

1-I liked particularly the special negro series of articles in a newspaper about further education and better housing.
2-No.
3-No.
4-No.
5-No.
6-
7-No.
8-

Group Y (str, nsp)

11-
12-Not against the races; just about the negroes.
13-No.
14-Not aside from here.
15-Book at school and the encylopedia.
16-No.
17-No
18-No.

CONCLUSION SITUATION I--IN "RACE" LEARNING THE SCALE-TAUGHT GROUP APPEARS MODERATELY SUPERIOR TO THE OTHER, PARTICULARLY IN THE QUALITY OF THE REPLIES, THOUGH BUT SLIGHTLY IN THE TABULATION (THREE CLASS INTERVAL SHIFTS):

"Race" Answers Situation I to Interview Section II Question Four

	Group Y (str)	Group X (nsr)
No	4	5
?	2	2
Yes	0	0
Definite Instances	2	1

CONCLUSION SITUATION II--NO MENTION IS MADE OF "PEACE" READING SO NO CONCLUSIONS CAN BE DRAWN.

Answers by Groups S and R, learners in
Situation III: Sstr, Rnsr, and in
Situation IV: Snsp, Rstp

Group S (str, nsp)

121-No.
122-Looked up the Negro in my social studies book at school but it didn't say much.
123-No.
124-
125-No.
126-No.
127-I don't know; I don't read much.
128-Yes, I read about the negroes.

Group R (nsr, stp)

21-
22-Read about drills.
23-
24-Looked for something, but couldn't find anything.
25-No, don't read much--have no time.
26-I don't know.
27-Didn't read anything.
28-No.

CONCLUSION SITUATION III--IN "RACE" LEARNING, THE SCALE-TAUGHT GROUP IS MODERATELY SUPERIOR, PARTICULARLY IN THE QUALITY OF THE ONE INCIDENT. THE SUPERIORITY IS SHOWN AS SLIGHT IN THE TABULATION (TWO CLASS INTERVAL SHIFTS):

"Race" Answers Situation III to Interview Section II Question Four

	Group S (str)	Group R (nsr)
No	5	4
?	1	3
Yes	0	1
Definite Instances	2	0

CONCLUSION SITUATION IV--IN "PEACE" LEARNING, THE SCALE-TAUGHT GROUP, GROUP R, SEEMS SLIGHTLY SUPERIOR TO THE OTHER, SINCE THE ONLY MENTION IN THE INTERVIEW OF READING ABOUT PEACE OR WAR WAS MADE IN THAT GROUP.

Value of Interview Index to Changed Action.--These data seem to be the most weighty of all in assigning definite teaching values to measurement scales. Their results are most difficult to explain as the result of any other cause than definite teaching value. Of course, the learners, especially in the direct questions, may have been led to answer what seemed to be wanted, definite action in life as a result of the teaching. But that same impulse acted in the case of scale-taught and non-scale-taught groups in the same way. And it is in the comparison of the results from the two types of teaching that the data are found; the total amount of learning, which may have been counterfeited to some extent, influences the conclusions only slightly. Of course, there is the possibility that the interviewer identified the members of the variously taught groups, and asked the questions so that the teaching values would appear for the scales. But it

must be remembered that each learner had been taught in one situation with the scales and in one situation without them; and that the learners were purposely mixed up to avoid such a possibility. And the possibility that so intricate an arrangement would be remembered by the interviewer and unconsciously acted on even while trying not to--this possibility seems very remote.

2. The Testimony of Statistical Measurements

For several reasons the statistical data were less reliable than could be hoped: not all statements were marked both times; the students were unfamiliar with tests of this sort, were tired with necessary preliminary testing, and consequently tended toward mistakes and carelessness. And the difficulty of translating opinions into numbers was ever-present.

Instruments Used.--In the race teaching situations, there was used for measurement the Scales Used in Teaching,[1] and also fifteen statements from the "Attitude Test on World Relations."[1] In the peace teaching situations, the Scale Used in Teaching was not used for measurement; here fifteen peace statements from Chave's test constituted the sole measurement.[1]

Calculations.--An attitude score before and after the teaching was calculated for each learner and for each statement. The changes in attitude of the scale-taught groups were compared with the changes in the non-scale-taught groups, on the basis of both persons and statements. Degrees of superiority of one group over the other was again indicated by the terms "slightly," "moderately," "strongly" and "very strongly." These terms were based, as described above,[2] on class interval shifts. However,

---

[1]Samples given above, pp. 7ff.

[2]Pp. 19, 20.

the tabulations from which the necessary class interval shifts were calculated were based upon mathematically equal class intervals rather than the more subjective class intervals necessary in the interview analysis.

Statistical Results Summarized.--For each of the two situations, Situations I and III, Race Teaching, the two tests with two tabulations apiece were used. And for the other two situations, II and IV, measurement was attempted with the Chave-Lobingier Test alone. Thus there were a total of twelve statistical tabulations or indices on the relative learning of scale-taught and non-scale-taught groups.[1] Out of these twelve, only one showed a scale-taught group inferior. Three showed no difference in learning between the two types of teaching, while the other eight showed the scale-taught group superior: one slightly, four moderately, and three very strongly. Table II shows these results in summary.[2]

Tables III to XIV give the statistical results in detail. Those tables (uneven numbers) showing the change in pupils give: first, the Pupil Number;[3] then the score in race friendliness (the higher the score, the less enmity) before the experiment and show the change. These changes are tabulated into appropriate class intervals, and a conclusion drawn from the comparison of the changes in the scale-taught groups with those in the non-scale-taught groups.

---

[1]Two tabulations of each of two tests in the two situations total 8 indices; two tabulations of one test in two situations total 4 indices--a grand total of 12 indices.

[2]P. 34.

[3]P. 12.

Those tables (even numbers) showing the change in statements give: first, the number of the statement; then, the number of persons increasing and those decreasing in race friendliness as shown by their marking that particular statement; and, finally, the sum of all the score changes on that particular statement--the algebraic sum, that is, plus score changes are added and the minus score changes are subtracted. These score change totals for each statement are tabulated and conclusions drawn as in the changes for individual pupils.

The tables show the various situations, tests, and indices in the following order:

Table	Situation	Test	Index (tab'n)
III	I Ystr, Xnsr	Chave-Lobingier	Pupil Change
IV	I " "	" "	Statement Change
V*	I " "	Scale Used in Teaching	Pupil "
VI*	I " "	Scale Used in Teaching	Statement "
VII*	II Xstp, Ynsp	Chave-Lobingier	Pupil "
VIII*	II " "	" "	Statement "
IX*	III Sstr, Rnsr	" "	Pupil "
X*	III " "	" "	Statement "
XI*	III " "	Scale Used in Teaching	Pupil "
XII*	III " "	Scale Used in Teaching	Statement "
XIII*	IV Rstp, Snsp	Chave-Lobingier	Pupil "
XIV*	IV " "	" "	Statement "

Table II, which summarizes results, and two of the other tables, one of each type, follow:

---

*These tables are not included in this treatment. Of course, each table has been completed, but to include all of them would assign more space to statistical results than their worth seems to justify.

TABLE II

THE TEACHING VALUE OF MEASUREMENT SCALES IN RACE AND PEACE TEACHING AS SHOWN BY STATISTICAL MEASUREMENT--SUPERIORITY OR INFERIORITY IN LEARNING OF SCALE-TAUGHT GROUPS AS COMPARED TO EQUIVALENT NON-SCALE-TAUGHT GROUPS AS SHOWN BY THEIR RESPONSES IN PAPER AND PENCIL TESTS BEFORE AND AFTER THE EXPERIMENT.

Statistical Test	Superiority or Inferiority in Learning of Scale-Taught Groups as compared to Equivalent Non-Scale-Taught Groups.	
	Groups X and Y	Groups R and S
	Situation I, Ystr, Xsnr.	Situation III, Sstr, Rnsr.
RACE: Chave-Lobingier Tabulation by persons	0 (III,0)**	Moderate (IX,5-8) Inferiority *
Tabulation by statements	0 (IV,-1)	0 (X,-2)*
RACE: Scale Used in Teaching Tabulation by persons	Moderate (V,7) Superiority #	Moderate (XI,4) Superiority*
Tabulation by statements	Very strong Superiority (VI,12)	Slight (XIV,3) Superiority*
	Situation II, Ynsp, Xstp.	Situation IV, Xnsp, Rstp.
PEACE: Chave-Lobingier Tabulation by persons	Very strong Superiority (VII, 11)	Moderate (XIII,6) Superiority
Tabulation by statements	Very strong Superiority (VIII, 20)	Moderate (XIV,5) Superiority

*Results in this situation were lessened by the home attitude of one learner. Her mother protested against the teaching, as recorded p. 46.

**The Roman numeral refers to the number of the Table; Arabic numbers refer to the number of class interval shifts of superiority in scale-taught groups.

#Class of Superiority: Negligible, 1 Class Interval Shift; Slight, 2,3,4; Moderate, 5,6,7; Strong, 8,9,10; Very strong, 11 up.

TABLE III

THE TEACHING VALUE OF MEASUREMENT SCALES AS SHOWN BY STATISTICAL MEASUREMENT--SUPERIORITY OR INFERIORITY IN LEARNING OF SCALE-TAUGHT GROUPS AS COMPARED TO EQUIVALENT NON-SCALE-TAUGHT GROUPS.

Instrument: Selected statements from the Chave-Lobingier Test (p. 5, N.1 above).
Change in the individual pupils considered.

Situation I: Ystr, Xnsr (nomenclature explained p. 4).

Group X (nsr)				Group Y (str)			
Pupil No.	Score Before	Score After	Change in Race Friend-liness	Change in Race Friend-liness	Score Before	Score After	Pupil No.
1	157*	126	-31	-50	100	50	11
2	88	110	22	40	60	100	12
3	113	85	-28	-20	53	33	13
4	70	10	-60	47	13	60	14
5	-7	40	47	21	-6	15	15
6	50	133	83	-40	113	73	16
7	100	100	0	0	20	20	17
8	46	133	87	140	53	193	18

*The score is in terms of race friendliness; i.e., the higher the score, the less prejudice, see p. 7.

TABLE III (cont'd)

Tabulation of
Individual Changes
into Classes
Class Interval=1/20 total possible spread=20*

Change			Members of Y (str)	Members of X (nsr)
Strong gain	(51	)	1	2
Moderate "	(31	50)	2	1
Slight "	(11	30)	1	1
No change	(10	-10)	1	1
Slight loss	(-11	-30)	1	1
Moderate "	(-31	-50)	2	1
Strong "	(-51	)	0	1

CONCLUSION SITUATION I PUPIL CHANGE INDEX.--The difference in the changes of the two groups as measured by this standared is negligible, since in Y one gain was medium rather than strong and one loss was medium rather than strong (total of 0 class interval shifts).

Note: This table, the first of its type, is offered as an example of method of treatment, not as a sample of the statistical results of this experiment. These statistical results are outlined in Table II. By chance, these first two tables show no difference in learning resulting from the two methods of teaching.

---

*These data help to prevent distortion of results by manipulation of the size of the class interval. The minimum score in this instrument is -200, the maximum 200, so that the change may vary 400 units--the total possible spread. The class interval is a conveniently-sized section of the total possible spread used to tabulate the raw score and thus aid in the summarization of their results. Here a convenient size for the class interval is 20, or 1/20 of the total possible spread.

TABLE IV

THE TEACHING VALUE OF MEASUREMENT SCALES AS SHOWN BY STATISTICAL MEASUREMENT--SUPERIORITY OR INFERIORITY IN LEARNING OF SCALE-TAUGHT GROUP AS COMPARED TO EQUIVALENT NON-SCALE-TAUGHT GROUP.

Instrument: Selected statements from Chave-Lobingier Test (p. 5).

Change in the test statements considered.

Situation I: Ystr, Xnsr.

Statement No.	Y (str)			X (nsr)		
	Persons Decreasing	Persons Increasing	Algebraic sum of score changes	Algebraic sum of score changes	Persons Decreasing	Persons Increasing
1*	2	1	-1	0	2	2
2	1	2	4	6	0	3
3	2	2	0	0	2	10
4	3	2	-5	4	1	4
5	1	3	1	-2	1	0
6	1	3	3	2	2	3
7	2	0	-3	1	1	2
8	1	2	3	5	1	2
9	2	3	1	3	1	3
10	1	3	5	-1	2	1
11	1	1	0	0	1	1
12	1	3	3	-3	1	1
13	2	0	-4	-3	2	0
14	2	1	-3	3	1	1
15	0	3	9	3	0	1

*Examples of these statements are given p. 5, N.1.

TABLE IV (cont'd)

Tabulation of
Statement Changes into Classes
Class Interval=1/8 total spread=4

Made from Algebraic Sum alone;
Persons considered only in extreme cases.

Change			Statements Marked by Y (str)	X (nsr)
Strong gain	11		0	0
Moderate "	7	10	1	0
Slight "	3	6	5	6
No change	2	-2	5	7
Slight loss	-3	-6	4	2
Moderate "	-7	-10	0	0
Strong "	-11		0	0

CONCLUSION SITUATION I STATEMENT CHANGE INDEX.--Y, the scale-taught group in this situation seems negligibly inferior to X, the non-scale-taught group (total of 1 class interval shift).

Note: This table, the first of its type, is offered as an example of method of treatment, not as a sample of the statistical results of this experiment. These statistical results are outlined in Table II. By chance, these first two tables show no difference in learning resulting from the two methods of teaching.

B. Where Teaching Values Arise

The preceding data on the teaching value of measurement scales seem incomplete without some attempt to locate the points at which these teaching values arise. Of what specific aid are measurement scales in teaching? At what stages in the learning process can measurement scales be used effectively? A separate experiment for each stage or for each possible specific value would have to be set up if that question were to be answered with scientific objectivity. And that is beyond the scope of this experiment. However, the descriptive data do give some light on the problem. The investigator and assisting teacher can describe the values which seem to them to be clear, with appropriate illustrations and cases from the teaching situations.

Summary.--The values in the teaching use of measurement scales indicated by the descriptive data are: first, arousing interest and furnishing motivation; second, stimulating discussion; third, helping in preventing indoctrination; fourth, establishing connections with the home; and fifth, more specifically in aiding in various stages in the learning process.

What Makes a Scale Valuable in Teaching.--These values are of course not present in every situation. What conditions then are necessary to maximum teaching value in measurement scales?

1. Value is greatest in types of experience which are comparatively less pressing problems to the learner himself. Of course, all teaching situations should be so worthful that they furnish their own motivation. But in fields such as World Citizenship, the youth often finds himself so much a stranger that motivation at the beginning is particularly necessary. In these situations the scales render valuable service in furnishing motivation and stimulating discussion. Of course, other values, such

as preventing indoctrination or establishing connection with the home, would be as great in other situations as in those of this type.

2. Scales used as teaching instruments should be prepared with that end in view. Many very efficient measuring devices would not serve well as teaching instruments because that end was not considered in their construction. Teaching values must be striven for.

3. A vivacity and journalistic selection of facts are necessary for motivation and stimulation.

4. An unbiased completeness for preventing indoctrination must be established by scientific process.[1]

5. A breadth of appeal and ease of marking is necessary for establishing contact with the home.

1. Motivation Assisted

One primary value of the measurement scale seems to lie in a capacity for arousing interest and furnishing motivation. Opportunity for this use of the scales, as of most of the other uses, was in the inclusion of the scales in the notebooks of each member of the scale-taught groups. Of course, all the learners at the beginning of the experiment marked the scales,[2] so that the difference between scale-taught and other groups was not so great as if one type of group were totally unfamiliar with the scales. This opening phase of the teaching use of scales may be seen by a comparison of the procedure of the teacher in arousing interest and furnishing motivation both with and without

---

[1]Procedure in perfecting the Scales Used in Teaching is described p. 7.

[2]With the exception of the peace unit situations; see p. 7, N.1.

scales in the first unit, "Is the race problem important enough to study?"

An Example: Teaching Without the Scales.--A realization of the situation, how some people consider the race problem of vast importance while others hold it not worth bothering about, and a definition of the issues involved were developed through discussion. The teacher helped to contribute to the discussion facts--and sometimes facts startling or paradoxical enough to induce the learner to see not only the bare outline of the situation but also its emotional hues. For example, at appropriate openings, the teacher remarked that Belgium's colonial empire is 88 times her own size, and quoted Achmet Abdullah, and W.E. Dubois on the war of the color line.[1] With these statements and facts as instruments, he endeavored to stimulate discussion, to bring out personal experiences which would lead to a realization of the situation, to inspire the group to search race sources. However, even though the facts, statements, and situations used by the teacher were identical with those contained on the scales, it was often difficult to produce them clearly or quickly enough to inspire and hold the interest.

With the Scales.--In the scales the teacher had a compact and elastic instrument with all these same statements, situations, and facts immediately available. The third section furnished a series of attractive paradoxes and facts important but outside the learner's range of experience. The pupil began marking the

---

[1] "'You (the white man) are deaf to the voice of reason and fairness, and so you must be taught with the whirring swish of the sword when it is red.' '...colored folk have much to remember and they will not forget...they are going to fight, and the War of the Color Line will outdo in savage inhumanity any war this world has yet seen.'" Lothrop Stoddard, The Rising Tide of Color (New York: Scribners, 1920), pp. 11,14.

scale probably through no more serious motive than an urge for guessing the right answer on Belgium's colonial empire, but as statement after statement appeared contrary to his preconceived notions he began to show an interest in the subject itself. He was startled and aroused by this simultaneous volley of thoughts, not merely disturbed by the occasional sniping shot possible in regular teaching tempo.

Discussion began almost automatically and a great deal of time was saved by allowing the pupils to choose from the scales the question most interesting to them. Interests could be located and stimulated more quickly and surely than by trial and error.

Without the scales, the teacher stimulated momentary flashes of interest by statements such as "It is better to eat a man without cooking him than to cook a man without eating him" (cannibalism versus lynching). But with the scales, he said simply, "Here are some question blanks that I think you will find interesting." And after a few directions, and a few minutes for marking, the discussion became so intense that it needed dampening, not further stimulation.

Section I of the scales helped to keep the interest from spending itself on some side track by listing various definite positions which may be taken on this question of the importance of the race problem.

Motivation Value in General.--Time after time, before this experiment was begun, the investigator has seen dull groups brighten and quicken at the touch of this simple device. And though the predispostion against tests caused in this experiment by the necessity of preliminary ones delayed this reaction, the quickening and stimulating came just as noticeably. Attempts toward the same end in the non-scale groups were much less

successful in promptness and intensity. The pupil may begin the paper through the "guessing urge." But almost immediately he finds himself in a strange "Alice Through the Looking Glass" world, where simple and understandable facts assail his assumptions, where everything seems just opposite to what he has always taken for granted. Does he believe the Negro an inferior being? Facts march past him, biological facts, anthropological facts, historical facts, quotations of experts, all so strange as to be almost unbelievable. And various other positions appear, all stated in attractive terms, with possibilities of action in definite situations which he had never considered before.

2. Discussion Stimulated

Closely allied with this value is that of stimulating discussion. The assisting teacher reports:

> "With the sheets (the scales), discussion is fairly general; the suggestions come from the pupils and the applications are presented naturally and more completely than if the leader must suggest all of them...
>
> "Without the sheets the instructor does most of the talking, suggesting, and drawing of applications...Less ground is covered with the sheets because there is more discussion, more outside data presented."[1]

At least one of the situations in the "What Would You Do?" section seemed to appeal to each one in the group, and though there was sometimes difficulty to decide which situation to discuss, there was rarely a necessity for further stimulation to discussion by the teacher.

3. Aid Against Indoctrination

One of the values that appeared was rather surprising to the investigator. In teaching situations of this kind before this experiment, he had received the impression that much of the learn-

---

[1]This statement was made without any suggestion from the investigator.

ing recorded by his experimental scales was the result of indoctrination. For the learners seemed more interested in being sure they had the "right answer," than in appreciating the issue and coming to a thoughtful decision on it. When the change was made from the true-false statements to the present form in the section "What Do You Think?" it was with the purpose of improving the measuring value. But it seems also to have changed the instruments, at least to some extent, from temptations to indoctrination to positive aids in guarding against it.

With the pre-experiment scales of True-False opinion statements, time and time again requests would come from learners, "What is the right answer?" in spite of repeated efforts by the teacher to show that there was no "right" answer. And not a single request of that sort came from the experimental groups. They seemed to realize that the opinion statements were merely attempts to show all the various possible opinions that one could hold on each matter.

The statement of the assisting teacher indicates a similar realization of this anti-indoctrination value:

> "Without the sheets, the conclusion is more forcefully directed by the leader; the pupils follow that line of thought, and are influenced by the leader's apparent point of view. With them, the opinion of each pupil plays an important part in formulating the opinion of the group, as he expresses his own reactions and knowledge of the subject as suggested by the possibilities before him."

The reactions of the learners in the interviews following the experiment lend strong support to the value of the scales in working against indoctrination. One learner (Number 11) even mentioned the scale value in "getting everybody's opinion" in the second free expression question. To the first question in the third section of the interview,[1] concerning how the learner

---

[1]Answers given in full p. 12.

felt about using these "question blanks," the answers seemed convincing. Enthusiasm for the scales was strong and almost unanimous, but the most significant feature of the answers was the repeated reference to this anti-indoctrination influence. Some of the statements were: "They help to give sides to problems because they have various degrees of questions at the top and bottom, that is, the top was one extreme and the bottom another." Or: "I think it helps you improve your thoughts on the subject; Part of it is on one side, part on the other." Or: "Worthwhile, not too much like school; shows both sides of the question, shows how to act." Or: "Fairly interesting, made me think; those blanks knew what I'd do, made me stop and think as I looked over the situation."

The frequency of the statement that the scales showed both sides of the problem seems, together with the teacher's statement, to indicate rather strongly that this anti-indoctrination value (promotion of clear thinking and tolerance) is present in teaching with this type of measurement scales.

4. Home Connections Established

One value that the scales do seem to possess is a possibility of establishing connections with the home. After the marking of the section "How Well Can You Guess?" the offer was extended to the learners that they could take home blank copies of this single sheet and try it out on the persons at home. About half the learners availed themselves of this privilege. And that at least one carried out that plan is shown by this incident:

> The teacher of the pair R and S happened to be calling upon the mother of learner No. 11. While talking about the sons, the mother said, "J--- brought home one of the papers the other night. I went over it with him and was quite surprised how liberal he was, considering his father's attitude toward the Negro. But then, I have known Mr.----(the father) to go out of his way to do a Negro a favor."

Another incident furnishes an illustration of the way in which conflict with the home can be smoothed out, or the smoothing process helped, by the teaching use of measurement scales:

> At a church meeting of another type, the mother of one of the learners approached the teacher of groups R and S with obvious purpose.
>
> Mother: I want to talk to you about what you are teaching on a Sunday morning.
>
> Teacher: Yes, Mrs.---.
>
> Mother: ----- (a learner) says that you teach that we should treat the niggers as good as white folks, that you say they are as good as white people if not better. She says you talk as if you thought they were better than we are.
>
> Teacher: We did talk about the Negroes, so that we could understand them and decide how a Christian should act toward them.
>
> Mother: Well, ----- says that she doesn't want to come if she is told that niggers are as good or better than we are. And you know that is a serious matter for a mother when her child doesn't want to come to a Sunday School because of what they teach.
>
> Teacher: Did you look over those papers, those blanks, with -----?
>
> Mother: No.
>
> Teacher: You look over them with her and I think then you will be able to straighten her out.[1]

And it seems quite conceivable that the scales may have had some influence either in changing opinion or at least eliminating the objection to the teaching, for the pupil did not remain away.

5. Values in Specific Learning Steps

All these values have been conducted with the different phases of the learning process which in the most complete analysis have been divided into thirteen steps.[2] But in certain of these steps the scales have seemed to possess particular value.

---

[1]Another lady entered the conversation at this point, and though the talk continued on the race problem, it no longer applied to scale teaching values in connection with the home.

[2]W.C. Bower, loc. cit.

Realizing the Situation.--In this first step the scale is of peculiar value in being able to inform the teacher just where the pupils stand on the matter. The scale may show that the matter seems to them very secondary, or that they know little about it, or that they are surprisingly informed on the matter and need little teaching. For example, the boys were found to be remarkably informed on modern military equipment and its destructiveness, and the teaching was adjusted accordingly. The scales could fill in statements of opinion of which the learner had previously been unconscious. More chance was offered for comprehensive discussion than would have been possible merely by chance questions.

Search of Experience.--Investigation of personal and racial experience, which always goes hand in hand with proper analysis of the situation, was facilitated greatly. The facts and startling statements of Section III of the scales "What Do You Think?" stimulated the learners into a thorough study of race experience much more quickly and easily than was possible in non-scale-taught groups. The memory of the learner in the search of his own experience was particularly helped by the situations in the section "What Would You Do?"

Two recorded incidents proffered out of the personal experience of the learner both happened in groups taught at the time with the use of the measurement scales:

> Learner No. 2, during a discussion of how Negroes feel about the treatment they receive, volunteered: "I was at a party one time at a house where a colored woman was a servant. And someone called her 'Nigger' and she was awful mad. When the time came to eat, she wasn't going to serve him because he called her that."

It is quite possible that without the similar situation in the section "What Would You Do?" he would not have recalled

this incident. Another incident from personal experience was volunteered by Learner No. 28, also at the time in a group taught with the use of the measurement scales:

> "Well, I think war is horrible. I just went out here where they were firing a few little guns and I couldn't stand it. Imagine what it must be like when they fire big ones all the time and you don't know when one of them will hit you."

Here the direct connection is not so apparent, but the full consideration of the horror of war in the scale may well have been the stimulus for the recall.

Analysis of the Situation.--When teaching without scales the teacher in this step of the learning process leads the learners to break up the situation into its various factors. For example, in the first unit, the class sought all reasons why they should study the race problem and all reasons why they should not. Of course, since teaching in a type of experience so unfamiliar to the learners as world fellowship, the teacher found difficulty in keeping from leading the pupils into the analysis he himself made. Naturally, he tried to marshall the facts of personal and race experience into a group-determined mold. But the mold often bore too many tell-tale marks of his own chisel.

When teaching with the scales, he had to aid him an instrument from which most bias had been eliminated by a long and careful process. Important opinions were represented in one or more different ways. The pupils themselves by using the scales were able to formulate analyses much more their own than was possible without the scales.

Possible Outcomes.--In leading the group toward identification and choice of outcomes the teacher when teaching without the scales tried to consider all reasonable solutions. However, in units such as "How Treat Persons of Other Races?" the possible

outcomes are tremendously complex. A course of action which the teacher may not consider at all may be the very one selected by members of the group. Obviously, the teacher needed situations in which he could identify the various possible outcomes and lead the learners to make thoughtful choice. And these situations he took from his own or race experience, or tried to draw from the learners--a difficult task since situations in which the learners themselves took part were scarce or very difficult to identify.

The scales provided various features to aid in this process. Just as the first section on statements of opinion was most valuable in realizing and analyzing the situation, so here Section II, the section of situations, was most valuable in identifying the possible outcomes. In identifying the outcomes for this same unit "How Treat Persons of Other Races?" the following possible courses were tentatively outlined: treat them entirely as equals, treat them as equivalent but different in minor respects, treat them as inferiors either by direct scorn and discrimination or through a lazy and over-bearing benevolence.

Then Situation 2 of this "What Would You Do?" section was used in the identification of these tentative outcomes. The learner was asked to imagine himself on a school committee arranging for a dance; and some of the committee are in favor of barring the Negroes. He can choose to speak in favor of admitting the Negroes, thus maintaining their essential equality. He can keep quiet, or he can propose a more tactful way of keeping out the Negroes (holding the dance at a place which excludes them), thus admitting that he considers the Negroes different from the whites, though not stating how different or why. Or he can support the exclusion of the Negroes, thus backing up the thesis of their substantial inferiority. Other situations in the section can be

used as "screens" to make clear the meaning of the phrases, to add other possibilities, to subtract unlikely or unimportant causes, and to make a more close analysis of each stand. Thus an approach can be made to a decision which is consistent and in fullest accord with the facts both secular and religious.

Experimentation and Habit Formation.--These are perhaps the most difficult steps of the teaching process and ones in which much teaching fails. The teacher without scales must have led the learners through the other steps with such enthusiasm and success that enough momentum is generated to carry them through the discouragement of these later stages. Besides the help from the carry-over of scale teaching from other steps, the scales did provide several definite helps in this stage of learning. The use of the scales in establishing connections with the home and in stimulating reading was of particular value in experimentation and habit formation.[1] The learners could discover objections which they had not considered; there were discovered situations of which they had never thought. By discussion with their families and friends they could accomplish some of the necessary experimentation. They may be forced to retraverse many previous steps. They may have advanced no further in race friendship, they may even have retrogressed, but their opinions could be sharpened, strengthened and made ready for intelligent action. In this discussion, too, they could gather suggestions for situations in which habit formation could begin. They may have decided that all races should receive equal treatment, but they must see situations in which they can act out that quality if the opinion is to become active. The problem situations of Section II offered types from which they could reproduce many situations for habit formation.

---

[1]Such use is described above, pp. 28ff. and 45ff.

Of course, actual habit formation cannot take place on paper. The scales could not accomplish actual formation of habits in race friendship or international good-will. But the scales thus prepared the way for the formation of these habits.

Use of the scales for measuring at the end of the study offered another advantage, though not strictly a teaching use of the scale. The teacher could discover the attitude of the individuals and plan for personal conference or unobtrusive follow-up teaching on points perhaps negligible for the whole group. Knowing where a pupil stood, the teacher could stimulate formation of habits at the stage at which the learner was able to begin.

Other situations showing habit-forming values in teaching with scales may have developed, but appreciational and activity phases had to be kept as equal as possible between the groups, lest they by chance influence results out of all proportion to the influence the scales might have upon them. In the notebook work of groups X and Y (gathering data on the best in other races), however, there is one additional circumstance which may bear upon the question. Group X, the group taught the race units without the teaching use of scales, did not even begin a group notebook; while Group Y, taught with the scales, did develop a very creditable notebook of pictures and clippings. However, it must be admitted that very probably the complications of adolescent love had much more to do with that phenomenon than did the teaching value of the scales, for the two in the group committee who together fashioned the notebook were of opposite sexes and at the time very attentive to each other.

A Summary of the Experiment and Results.--"There is a profound conviction on the part of many students of education, that the effect of the testing movement has been to mechanize

education, and to twist and distort character." That statement came from a pioneer and leader in religious education for thirty years.[1]

The results of this experiment may not alter that conviction. But this experiment does show at least one situation in which measuring instruments have not acted to "twist and distort character," but have promoted an intelligent grasp of the situation, stimulated a careful study of experience, guided to a thoughtful choice of outcomes in a most complex situation--in short, have resulted in an increase of learning over similar teaching without the instruments.

The problem: What teaching value, if any, is there in measurement scales?

The instruments: Types of world citizenship experience were arranged as teaching units or plans. Measuring scales or tests were constructed for both teaching and measuring use.

The results: Measurement scales were proven to have definite teaching value, at least when used under similar conditions. In each of the four situations the scale-taught group displayed superior learning to the non-scale-taught group. This was shown by interviews with all learners, and by statistical results from tests taken before and after the teaching. (Statistical results on Situation III--Sstr, Rnsr--furnished the only exception. Here the learning in both groups was shown as approximately the same. But a reason for this exception was located.)

Another less objective result was the location of where this teaching value arises. Evidence was offered to show value

---

[1]G.A. Coe, in an unpublished lecture at the Divinity School of the University of Chicago, June 11, 1931. Whether such was his own conviction, he did not state.

in arousing interest and furnishing motivation, in stimulating discussion, in preventing indoctrination, and in establishing connections with the home. Value was also located at various stages in the learning process, in realizing the situation, in the search of experience, race and personal, in analyzing the situation, in locating and choosing possible outcomes, and in aiding in experimentation and habit formation.